AENEAS Project Overview

Michiel van Haarlem

AENEAS Plenary Meeting, Utrecht, NL

11-14 November 2019

Advanced European Network of E-infrastructures for Astronomy with the SKA

Design and specification of a distributed, European SKA Regional Centre to support the astronomical community in achieving the scientific goals of the SKA

EC Horizon 2020 (€3 million)

13 countries, 28 partners, SKAO, host countries, e-infrastructures (EGI, GÉANT, RDA), NREN's

Three year project (2017-2019)

- WP1: Project Management
- WP2: Governance Structure and Business Models
- WP3: Computing and Processing Requirements
- WP4: Data Transport and Optimal European Storage Topologies
- WP5: Data Access and Knowledge Creation
- WP6: User Services

Open Questions

Where will the SKA science archive data be hosted?

How will that data be transported from the sites to Europe?

How can we take optimal advantage of existing infrastructure?

What are the processing requirements and technologies to consider?

What interfaces, tools, and techniques will users need for analysis?

How do we setup and operate an international network of SRCs?

Advanced European Network of E-infrastructures for Astronomy with the SKA

Design and specification of a distributed, European SKA Regional Centre to support the astronomical community in achieving the scientific goals of the SKA

EC Horizon 2020 (€3 million)

13 countries, 28 partners, SKAO, host countries, e-infrastructures (EGI, GÉANT, RDA), NREN's

Three year project (2017-2019)

- WP1: Project Management
- WP2: Governance Structure and Business Models
- WP3: Computing and Processing Requirements
- WP4: Data Transport and Optimal European Storage Topologies
- WP5: Data Access and Knowledge Creation
- WP6: User Services

Previous Meetings

- 28 February - 1 March 2017 - NWO HQ, The Hague, NL
- 18-20 October 2017 - IAA, Granada, ES
- 26-28 March 2018 - OCA, Nice, FR
- 8-10 October 2018 - INAF, Bologna, IT
- 5-7 March 2019 - U Man, Manchester, UK
- 11-14 November 2019 - Utrecht, NL

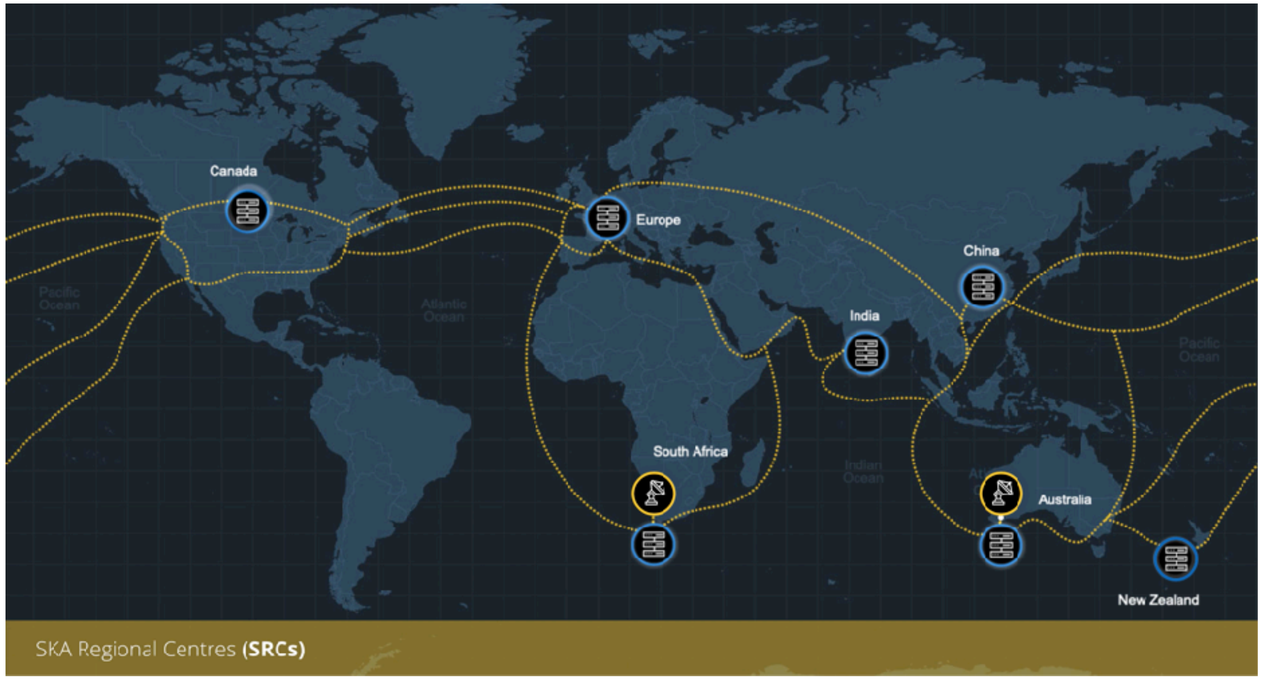
AENEAS results

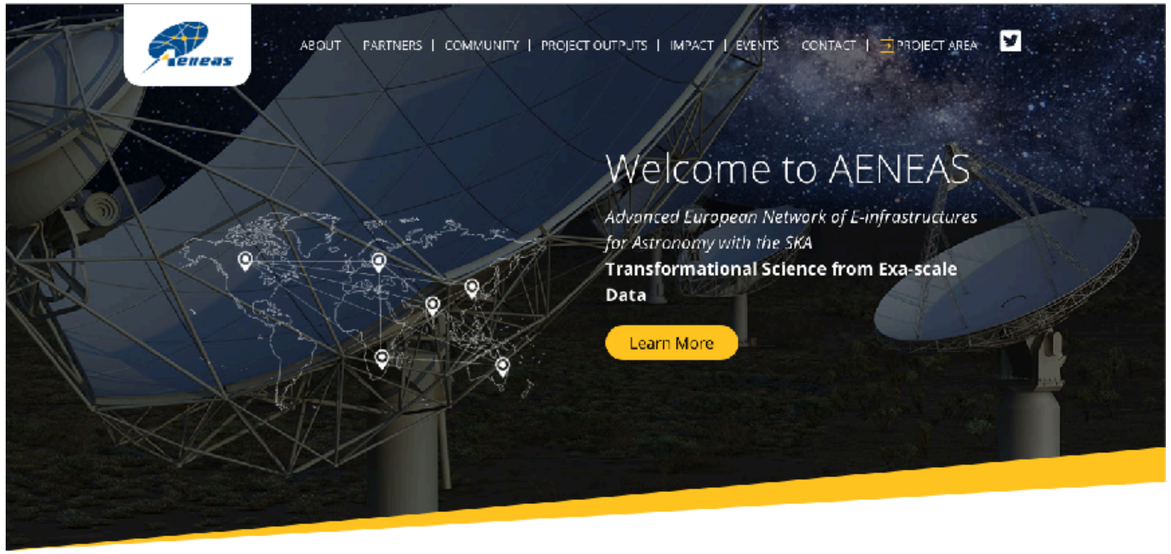
Available through web site: www.aeneas2020.eu

Key Project Findings

Deliverables & Milestones

Presentations





News & Announcements

A new treaty paves the way forward for the Square Kilometre Array

Mar 13, 2019 | News

Representatives from the founding member states of the Square Kilometre Array gathered in Rome yesterday to sign a treaty establishing the SKA Observatory as an intergovernmental organization that will oversee the delivery and operation of the world's largest radio...

read more

A step closer to a comprehensive design for the European SKA Regional Center

Mar 12, 2019 | News

The AENEAS team gathered at the University of Manchester last week for its 4th all-hands meeting. In addition to the usual project updates and presentations on regional center activities beyond Europe, much of the meeting was dedicated to focused discussions that...

read more

AENEAS and SKA collaborators convene in Manchester, UK

Feb 15, 2019 | News

The AENEAS team is convening in Manchester, UK for the 4th all-hands meeting on March 5-7. Along with the usual updates from team members, contributed talks by colleagues at the SKA office and partners from other SKA regional center, the meeting will...

read more

Final Plenary Meeting

- Final Review 4 March 2020 in Luxembourg
- EC programme officer, AENEAS MT, and external reviewers
- Final Deliverables due by end of 2019
- Final Report due for Final Review
- SRC activities underway in all SKA Member Countries
- SKA Regional Centre Coordination Committee (RSCSC) active
- Planning for future collaboration

Goals for this Meeting

- Provide overview of SRC landscape - in Europe and beyond
- Present and summarise AENEAS work
- Prepare for completion of project and final review
- Discuss future steps - technology, data challenges, governance
- Start planning for integration activity and implementation phase

WP3: Computing and Processing Requirements

Anna Scaife - Manchester University

Mark Ashdown - Cambridge University

Work package number	3	**Lead beneficiary**			UMAN
Work package title	Computing Requirements				
Participant number	1	2	3	4	5	7
Short name of participant	ASTRON	UMAN	UCAM	INAF	Chalmers	EGI.eu
Person/months per participant:	11	12	11	15	8	4
Participant number	9	11	12	13	20	21
Short name of participant	Jülich	STFC	CSIC	IT	EPFL	UNIGE
Person/months per participant:	9	6	3	6	3	3

Data Storage Requirements

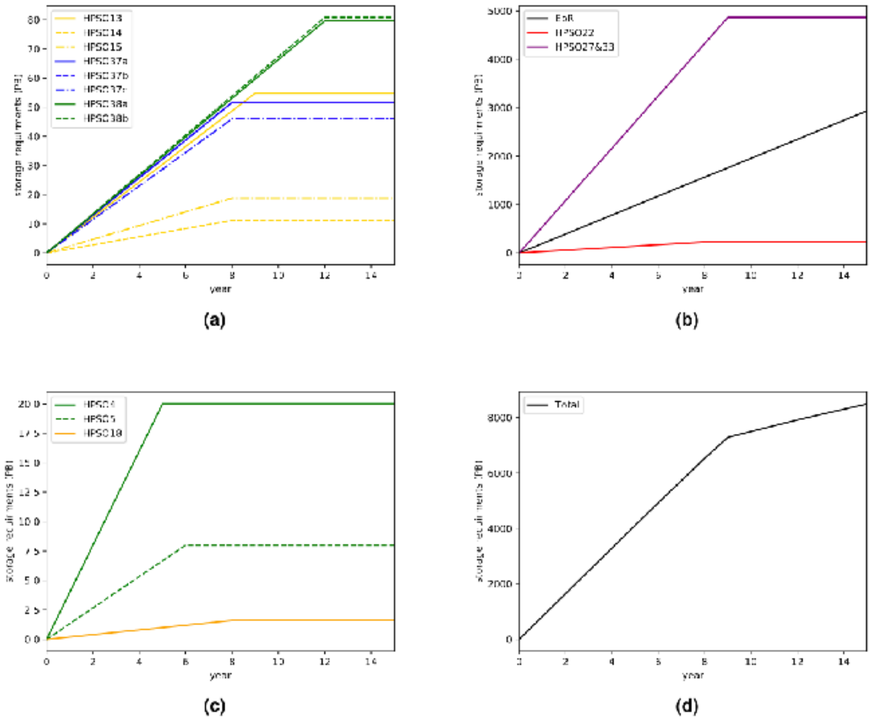


Figure 1: (a) Data storage requirements at SRCs for the HI and continuum HPSO. (b) Data storage requirements at SRCs for the EoR, magnetism and cradle of life HPSOs. (c) Data storage requirements at SRCs for the pulsars and transients HPSOs. (d) Data storage requirements at SRCs for all HPSOs.

Minimum storage needed to meet the needs of the HPSOs:

- **744 PB per year for the first 10 years of operation**, and
- **201 PB per year for the following 5 years**

The HPSO storage volume required by the SRC network is therefore **8.5 Exabytes over the course of 15 years**

Rate of production of advanced data products was determined using the AENEAS processing Use Cases and was determined to be **3:1 in volume** (output:input)

Also examined the **number** of data products expected as a function of time. Relevant for the choice of Data Management System (DMS)

Data Processing Requirements

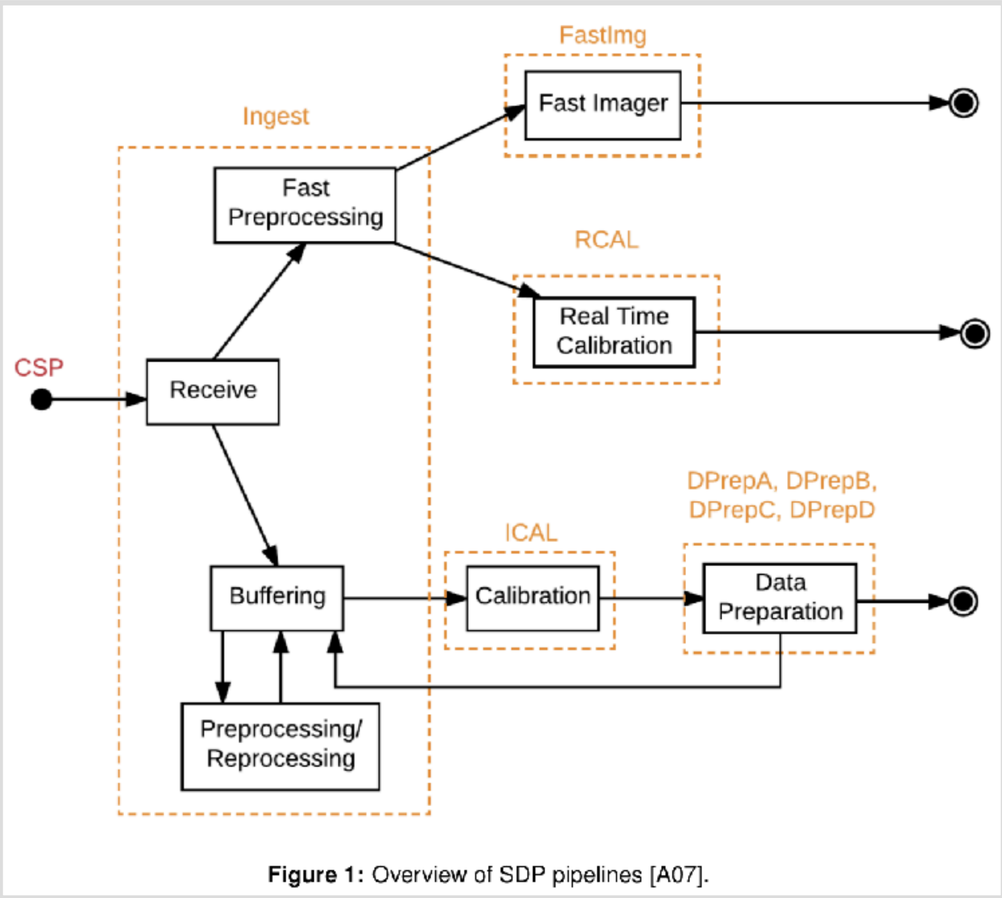


Figure 1: Overview of SDP pipelines [A07].

Minimum processing required to meet the needs of the HPSOs: **~26 PFlops**

This is subject to **strong** assumptions:

- Re-processing / post-processing is performed only **once** for each primary data product;
- That processing of SDP data products for all HPSOs is performed **at the same rate that they are ingested** into the SRCs

Software Infrastructure

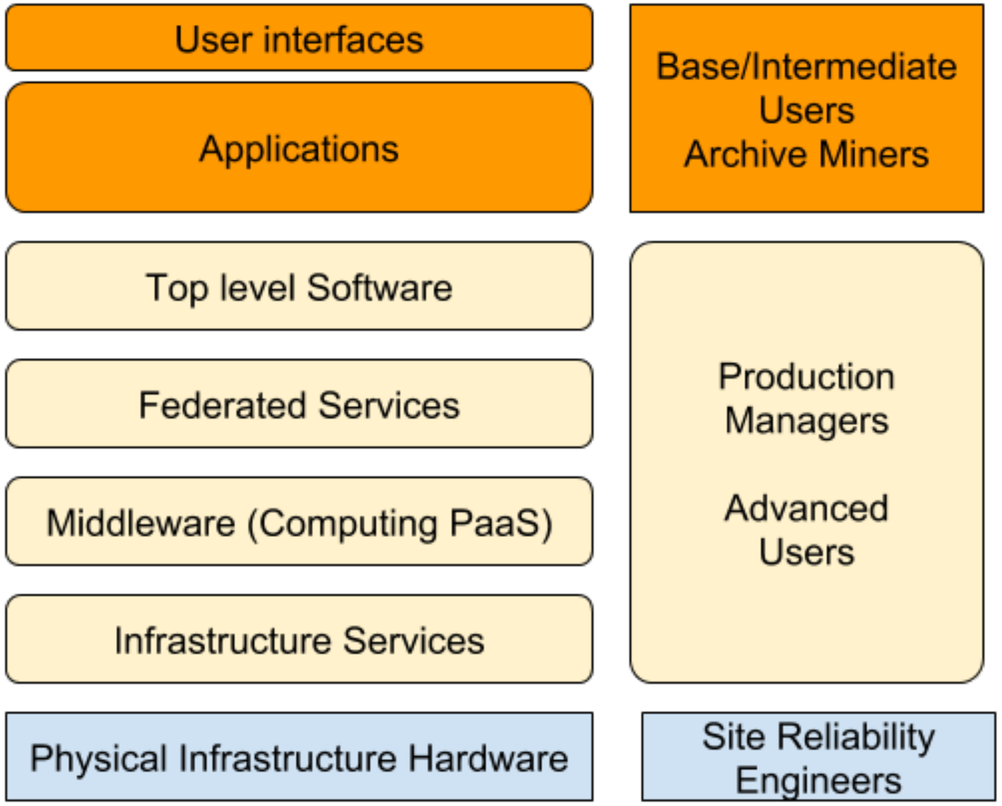


Figure 2.1: ESDC software stack and users' interaction.

- **Top-level Software** layer comprises all software components users interact with to model and run their distributed applications
- **Federated Services** layer consists of all the components expected to be provided by each site to support SRC federation
- **Middleware (or Computing/PaaS)** layer is the set of orchestration and support services necessary to support distributed compute workflows within SRC infrastructure
- **Infrastructure Service** layer includes basic services for setting up and maintaining a computing site infrastructure

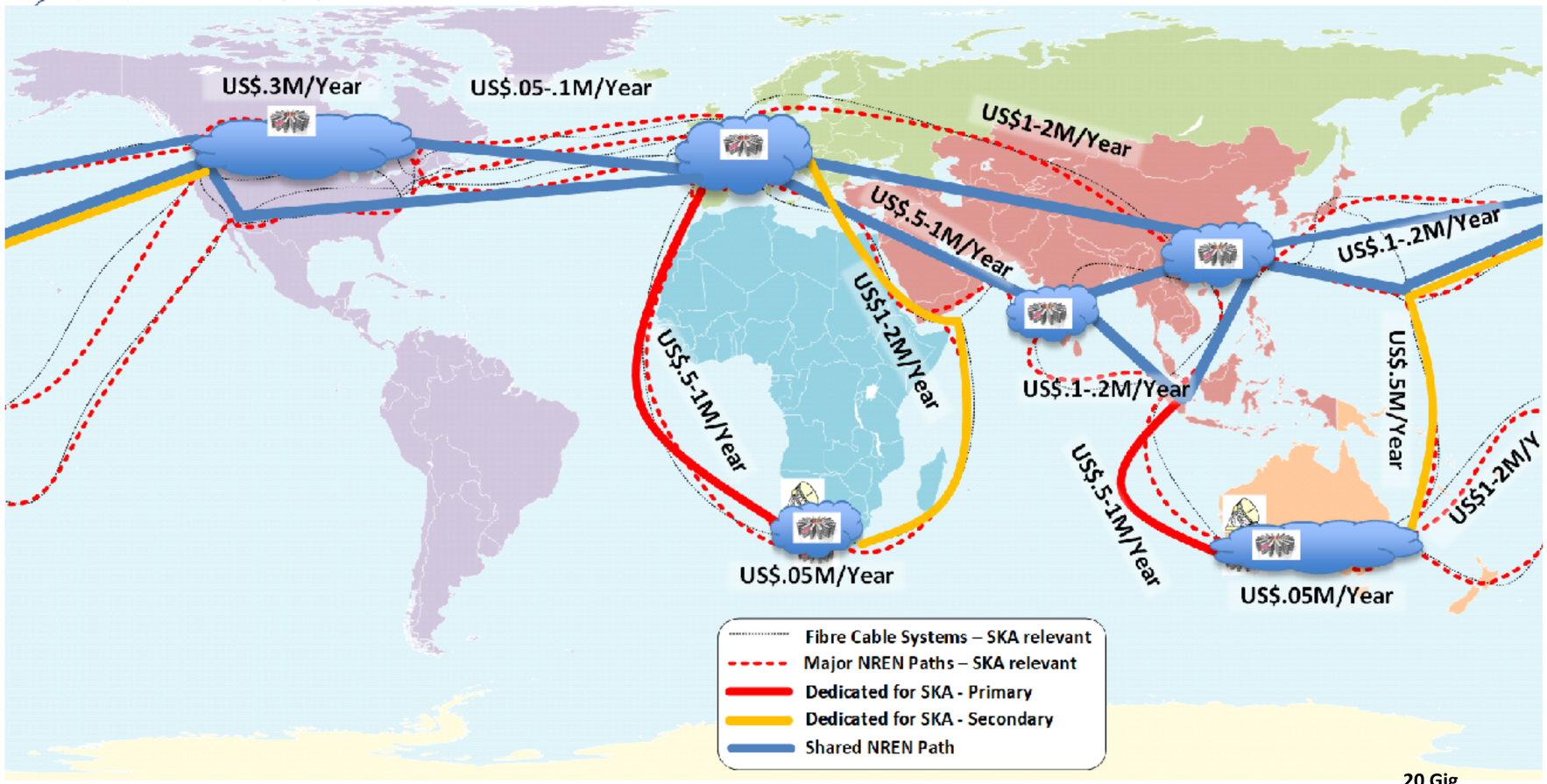
Identified four categories of users: **basic**, **intermediate**, **advanced**, and **production managers**, and made an assessment of their requirements

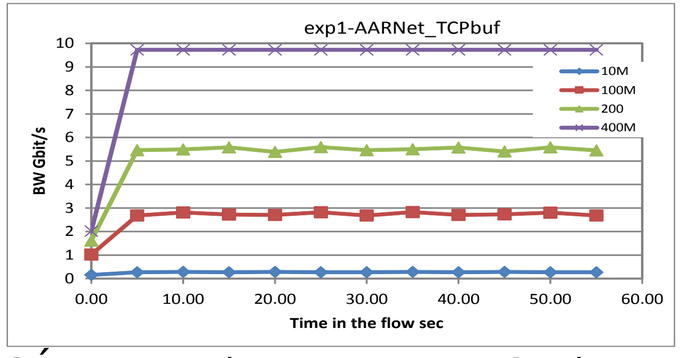
WP4: Data Transport and Optimal European Storage Topologies

Richard Hughes-Jones
GEANT

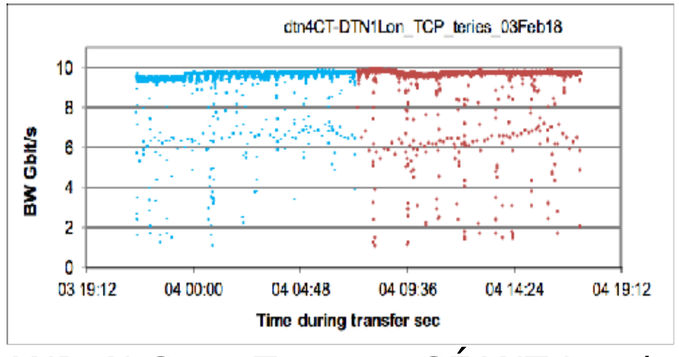
Work package number	4	**Lead beneficiary**		GEANT LTD	
Work package title	Analysis of Global SKA Data Transport and Optimal European Storage Topologies				
Participant number	2	4	5	6	8
Short name of participant	UMAN	INAF	Chalmers	GEANT LTD	Jülich
Person/months per participant:	6	1	20	22	9
Start month	1		End month	36	

WP4: Data Transport and Optimal Storage Topologies



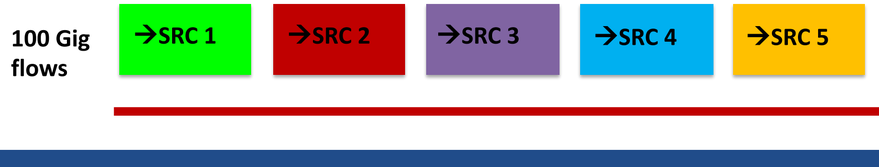


GÉANT London to AARNet Canberra



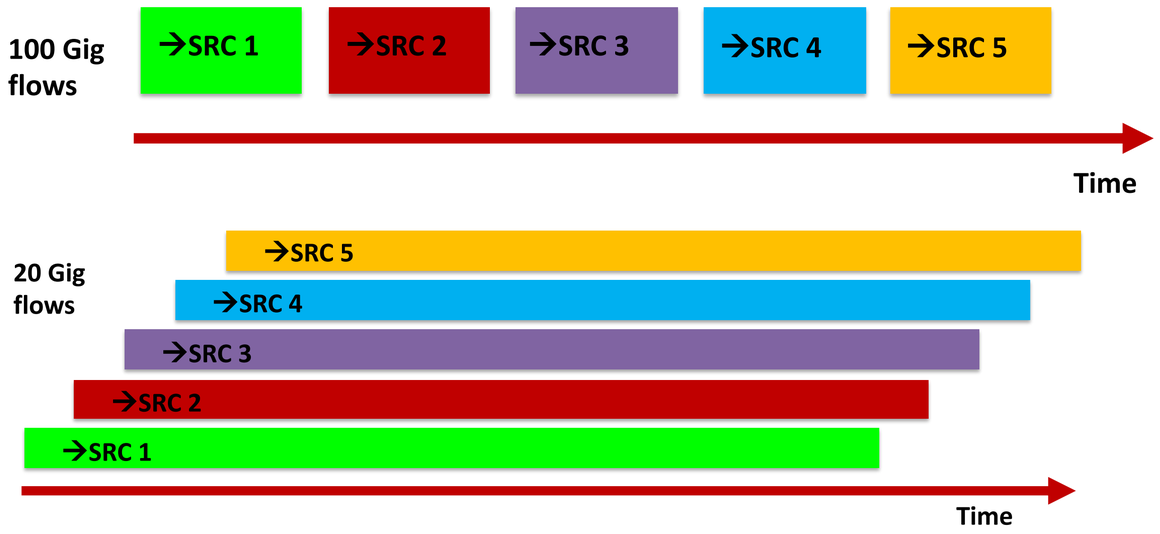
SANReN Cape Town to GÉANT London

What fraction of the SKA archive will we host in Europe?

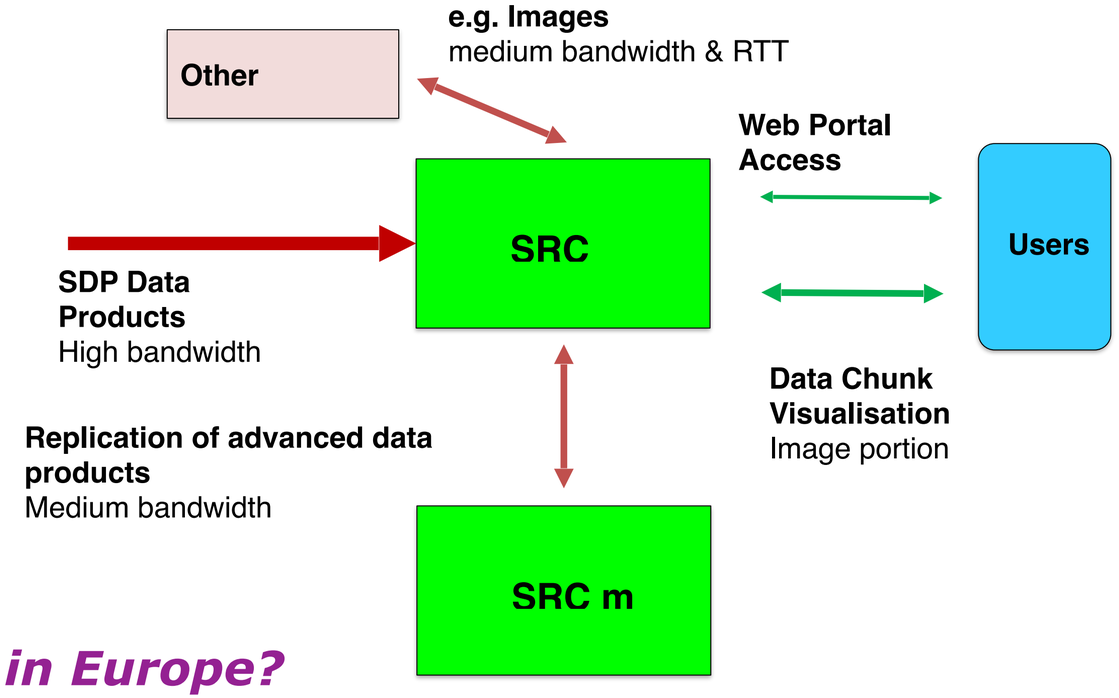


Models of Global SRC Data Flow

Distribution from Sites to SRC



Interactions between SRCs and



Which fraction of the SKA archive will we host in Europe?

WP5: Data Access and Knowledge Creation

Marcella Massardi

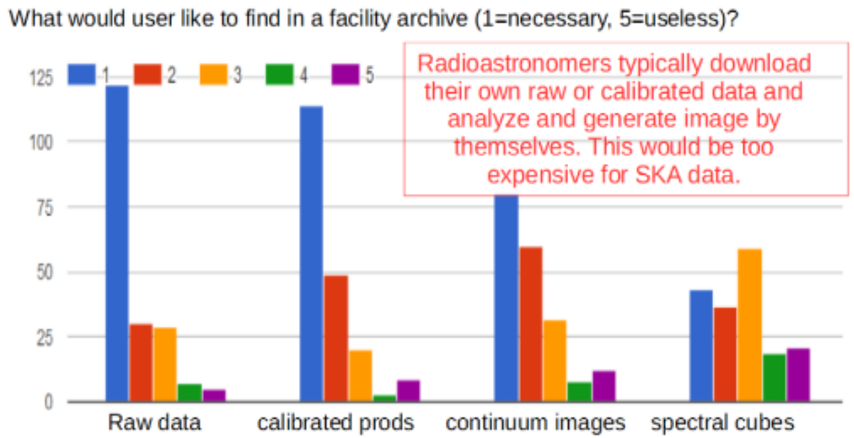
INAF

Work package number		**Lead beneficiary**			INAF	
Work package title	Access and Knowledge Creation					
Participant number	1	2	3	4	12	14
Short name of participant	ASTRON	UMAN	UCAM	INAF	CSIC	CNRS
Person/months per participant:	12	6	6	27	6	6
Start month	1		**End month**	36		

AENEAS WP5

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016

M0 M12 M18 M24 M34

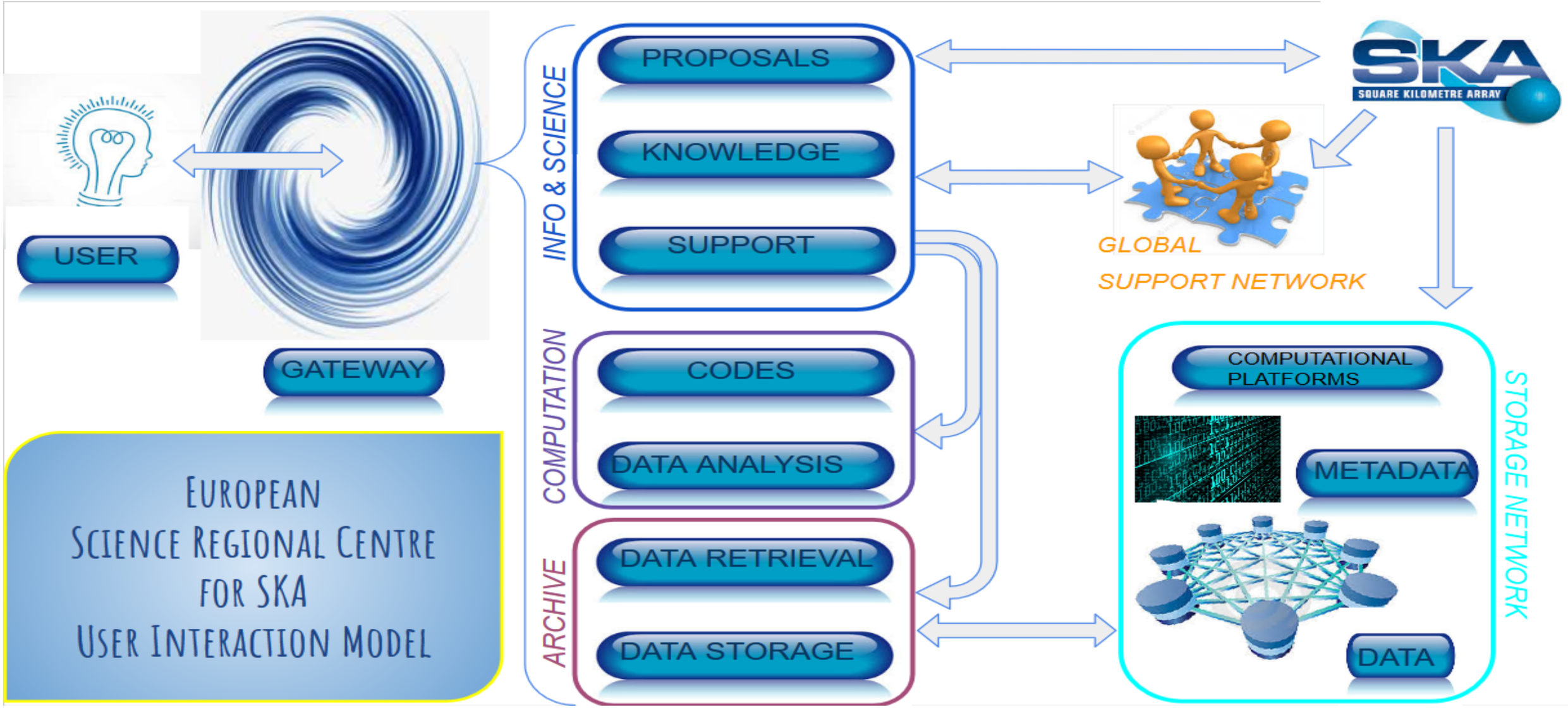


Learning from the past...

1) System needs
(goals towards user&tel)
2) User definition
(community/mentality)
3) Services provided
(duties/activities/policy/limitations)
4) Accessibility
(human interaction/interfaces)
5) Resources
(personnel/tools/infrastructures)

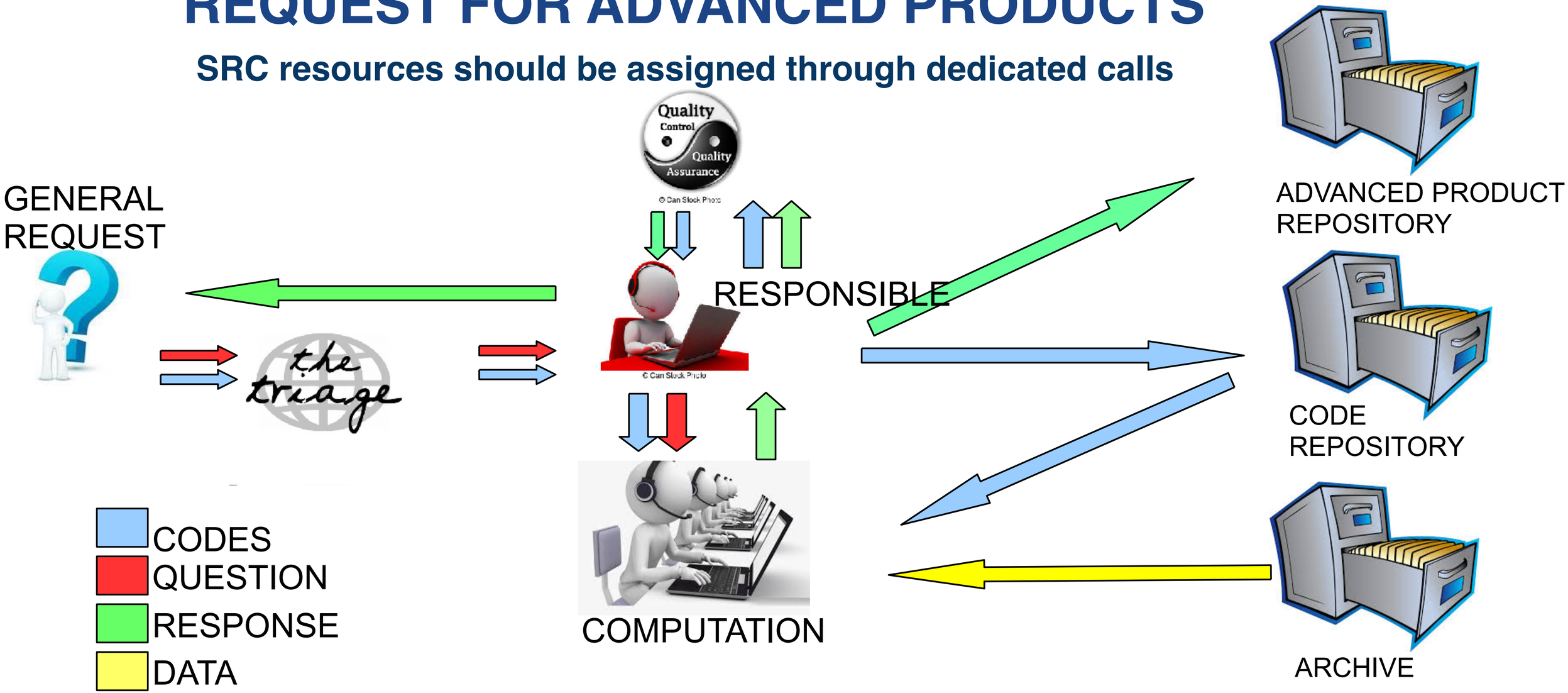
The ESRC is the interface, access the archive, offers the computation platform. It must be trustworthy and resilient. An efficient Regional Centre IS A RESOURCE FOR THE USERS

...suggesting new ideas


USER
GATEWAY
INFO & SCIENCE
PROPOSALS
KNOWLEDGE
SUPPORT
COMPUTATION
CODES
DATA ANALYSIS
ARCHIVE
DATA RETRIEVAL
DATA STORAGE
SKA
SQUARE KILOMETRE ARRAY
GLOBAL
SUPPORT NETWORK
COMPUTATIONAL
PLATFORMS
METADATA
DATA
STORAGE NETWORK
EUROPEAN
SCIENCE REGIONAL CENTRE
FOR SKA
USER INTERACTION MODEL

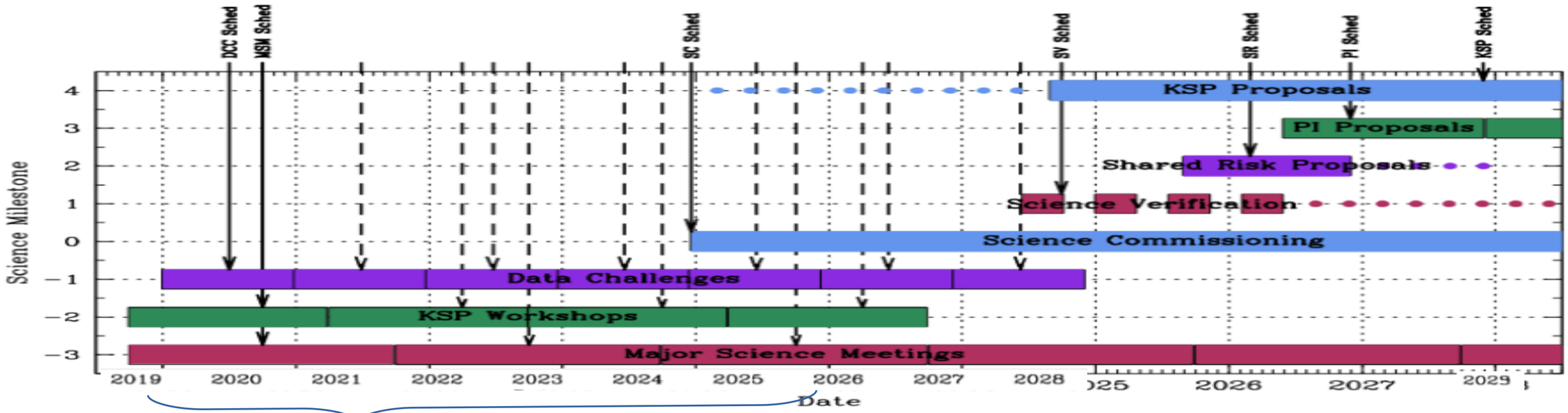
REQUEST FOR ADVANCED PRODUCTS

SRC resources should be assigned through dedicated calls



A training plan for the SKA community

Advanced European Network of E-infrastructures for Astronomy with the SKA AENEAS - 731016



Training phases:

Cognitive

-acquiring the general overview and basic knowledge on the system capabilities (i.e. data structure and handling)

-building awareness on the gaps and issues

Practice

- dirty hands-on simulations, data from precursors and lately on SKA data
- test tools/interfaces and workflows

Active

- proposing as PIs
- mining the archive
- maintaining and distributing the knowledge

WP6: Services

Matthew Viljoen
EGI Foundation

Work package number	1	**Lead beneficiary**		EGI.eu	
Work package title	Services				
Participant number	4	6	7	15	16
Short name of participant	INAF	GEANT LTD	EGI.eu	GRNET	FOM
Person/months per participant:	3	4	16	6	6
Start month	1		End month	36	

WP6 Objectives

Tasks	Project Objectives
T6.1 Federated Authentication, Authorization Infrastructure (AAI) and Identity Provisioning (AAI) *(Lead: **INAF**)*	• Collection of AAI Requirements • Recommendation of approaches and solutions • Proposal of a trust model
T6.2 Interoperabile Federated IT Service Management (ITSM) System *(Lead: **EGI Foundation**)*	• Assessment of existing structures and tools at national level and the relationship with SKA SRCs • Analysis of applicable standards/approaches of federated service management • Recommendation for an operational architecture and core FitSM processes
T6.3 Federated ITSM Support Tools *(Lead: **EGI Foundation**)*	• Recommendation of tools to support Federated ITSM in the ESDC Network • Piloting of an integrated system of ITSM support tools in different e-Infrastructures supporting the ESDC network

T6.1 Federated AAI

- Collected AAI requirements from ESRC using input from different parts of project
- Recommended approach based on AARC Blueprint Architecture and trust model (D6.1)
- Ran a series of piloting activities
  - integration of collaboration tools with FedAAI
  - integration of community IdP with FedAAI
  - working with tools providers (e.g. Dirac) testing with RCauth & providing community requirements

T6.2 Interoperable Federated ITSM

- Proposed framework for designing and implementing a Service Portfolio for the ESDC and SKA (D6.2)
- Set up a sample ITSM framework for the project
- Recommendations for ITSM support processes (D6.4)

T6.3 Federated ITSM Support Tools

- Evaluation of suitable ITSM support tools including collaboration tools
- Lessons learned from setting up EOSC ‘federation of federation’ (EOSC-hub project)
- Final comprehensive recommendations of support tools and operational support structures suitable for federated ESDC (D6.4)

WP2: Governance Structure and Business Models

Michiel van Haarlem - ASTRON

John Conway - Onsala

Work package number	2		**Lead beneficiary**			ASTRON
Work package title	Development of ESDC Governance Structure and Business Models					
Participant number	1	2	3	4	5	6
Short name of participant	ASTRON	UMAN	UCAM	INAF	CHALMERS	GEANT LTD
Person/months per participant:	12	1	1	2	2	2
Participant number	7	28				
Short name of participant	EGI.eu	RDA				
Person/months per participant:	2	2				
Start month	1			End month	36	

WP2: ESDC Design & Governance

Survey of Potential Providers

- *Over 50 expressions of interest*
- *Mixture of scientific institutes, infrastructure providers, and industrial partners*

- *ESDC Requirements based on those developed by SRCCG*
- *Final deliverable: preliminary ESDC Design and Implementation Plan*
- *User input needed!*

Strategic Partnerships and Next Steps

CIRADA
CANADIAN INITIATIVE FOR RADIO ASTRONOMY DATA ANALYSIS

ERIDANUS
Exascale Research
Infrastructure
For Data In Asia-Pacific
Astronomy
Using The SKA

CERN-SKA MoU

ESCAPE
European Science Cluster of Astronomy & Particle physics
ESFRI research infrastructures

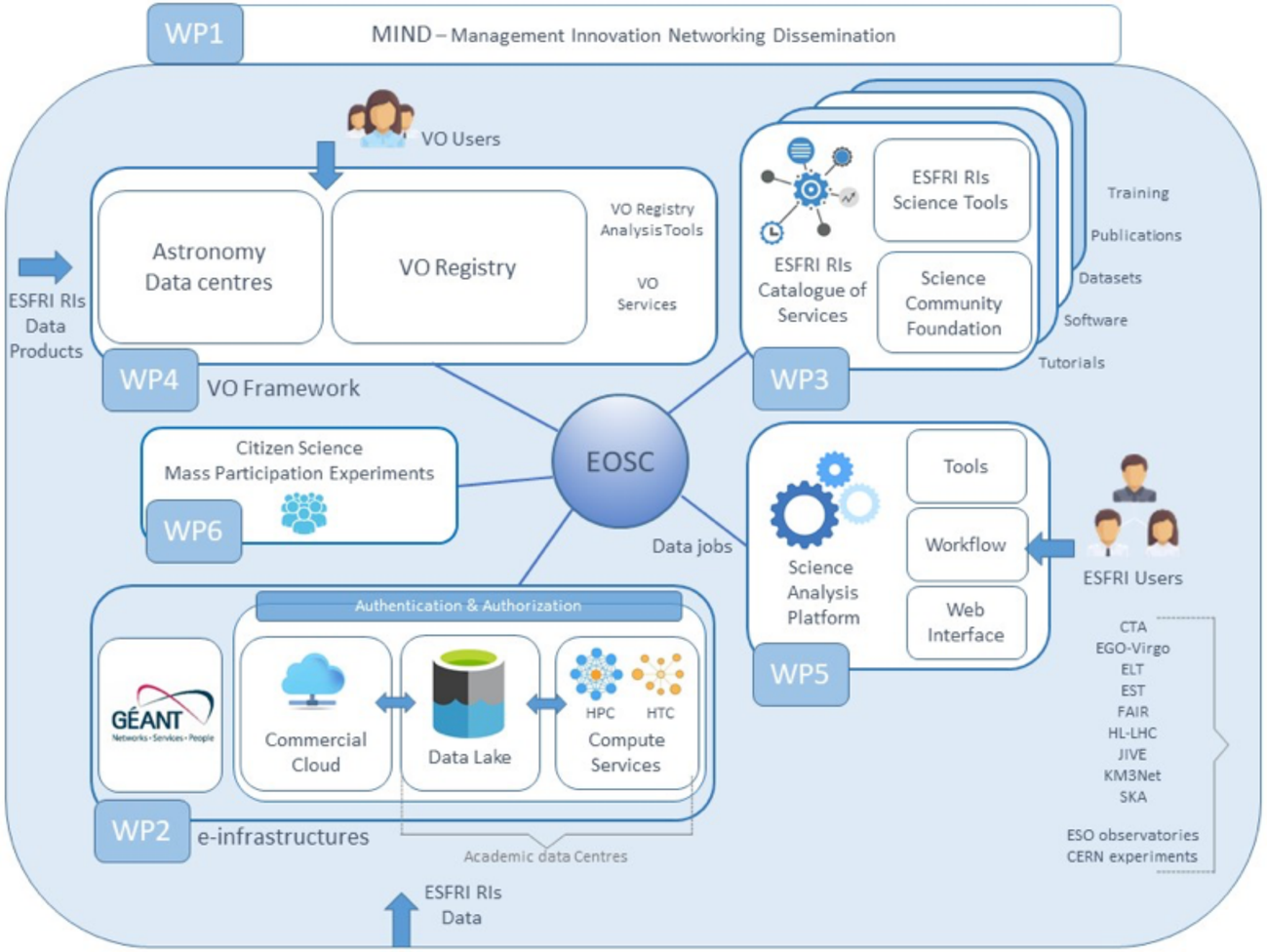
EOSCpilot
The European Open Science
Cloud for Research Pilot Project

EOSC-hub

ESCAPE

European Science Cluster of Astronomy & Particle physics
ESFRI research infrastructures

ESCAPE aims to address the Open Science challenges shared by ESFRI facilities as well as other pan-European research infrastructures in astronomy and particle physics



- EC H2020 (16 M€, 2019-2023)
- Partners include SKA, CTA, KM3Net, EST, ELT, HL-LHC, FAIR, CERN, ESO, JIVE
- Led by CNRS, 32 different EU institutions
- ASTRON leading Science Platform WP
- Work kicks off in January 2019

Goals for this Meeting

- Provide overview of SRC landscape - in Europe and beyond
- Present and summarise AENEAS work
- Prepare for completion of project and final review
- Discuss future steps - technology, data challenges, governance
- Start planning for integration activity and implementation phase